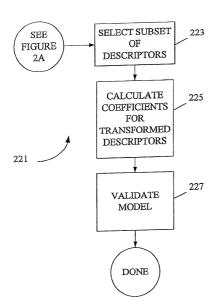


FIGURE 2A



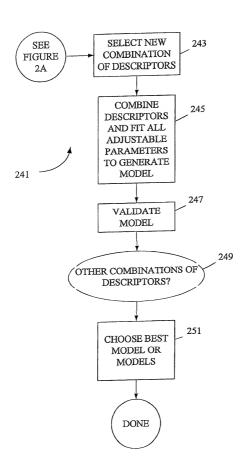
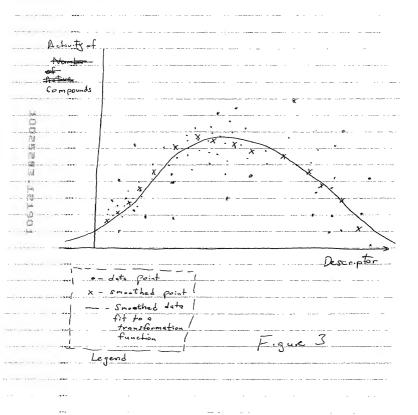
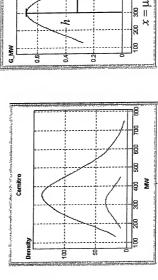
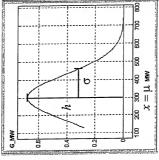


FIGURE 2C



## Optimum Molecular Weight





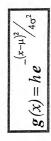
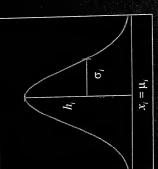


Figure 4A

# N-Dimensional Gaussian Modeling

#### Additive

$$m{g}(x_{i..N}) = rac{1}{N} \sum_{i=1}^{N} h_i \, m{e}^{-(x_i - \mu_i)^2 / 4\sigma_i^2}$$



#### Multiplicative

$$g(x_{i.N}) = h e^{-\frac{1}{N} \sum_{i=1}^{N} (x_i - \mu_i)^2 / 4\sigma_i^2}$$



Figure 4 B

### **Optimization Function**

$$\mathbf{g}\left(\mathbf{x},\boldsymbol{\mu},\boldsymbol{\sigma},\boldsymbol{h},t\right)=t+h\,e^{-\sum_{k=1}^{N_{x}}\left(x_{k}-\mu_{k}\right)^{2}/4\sigma_{k}^{2}}$$

$$f = s_{puh} \left[ \frac{1}{N_{mh}} \sum_{i=1}^{N_{mh}} \left( \mathbf{g}(\mathbf{X}_i, \boldsymbol{\mu}, \boldsymbol{\sigma}, h, t) - \mathcal{Y}_i \right)^2 
ight]$$

$$+ S_{drig} \left[ \frac{1}{N_{drig}} \sum_{j=1}^{N_{drig}} \mathbf{g}(\mathbf{X}_j, \boldsymbol{\mu}, \boldsymbol{\sigma}, h, t) - \overline{y}_{drig} \right]^2$$

Mean of the Squared Errors of Inhibitor Affinity Squared Error of the Means of Drug Affinity

Constraints to prevent Overfitting

 $+ s_{fi} \left| \sigma_y^2 \sum_{k=1}^{N_t} \left( \frac{\mu_k - \mu_{0,k}}{range(X_k^T)} \right)^2 + (t - t_0)^2 \right|$ 





## Initial Values for Optimization

$$t_{0} = min(y)$$

$$h_{0} = max(y) - t_{0}$$

$$\sum_{i=1}^{N_{max}} (y_{i} - t_{0})^{2} x_{k,i}$$

$$\mu_{0,k} = \frac{\sum_{i=1}^{N_{max}} (y_{i} - t_{0})^{2} x_{k,i}}{\sum_{i=1}^{N_{max}} (y_{i} - t_{0})^{2}}$$

$$\sigma_{0,k} = \sqrt{\sum_{i=1}^{N_{max}} (y_{i} - t_{0})^{2}}$$

$$\sum_{i=1}^{N_{max}} (y_{i} - \overline{y}_{m,h})^{2}$$

$$\sigma_{y} = \sqrt{\sum_{i=1}^{N_{max}} (y_{i} - \overline{y}_{m,h})^{2}}$$

omikro S

# Gaussian Optimization Function

$$f(\mathbf{x}) = t + h e^{-\sum_{k=1}^{N_k} (x_k - c_k)^2 / 4w_k^2}$$

$$f(\mathbf{x}) = \int_{\mathbf{x}} \int_{\mathbf{x}=1}^{N_{v}} u_{i} (f(\mathbf{x}_{i}) - y_{i})^{2}$$

$$= s_{y} \left( \int_{\mathbf{x}=1}^{N_{v}} u_{i} (f(\mathbf{x}_{i}) - y_{i})^{2} \right)$$

$$S_c \sum_{k=1}^{N_c} \frac{\left(c_k - c_{0,k}\right)^2}{\sigma_{s_k}^2}$$

$$S_{W} \sum_{k=1}^{N_{X}} \frac{\sigma_{X_{k}}}{W_{k}}$$

$$S_t \frac{(t-t_0)}{\sigma_y^2}$$

Weighted Mean Squared Error

#### Center Constraint

Width (Focus) Constraint

Tare Constraint



### Gaussian Optimization Starting Values

$$\sum_{i=1}^{N_{math}} u_i (y_i - \overline{y})^2$$

$$\sum_{i=1}^{N_{math}} u_i (y_i - \overline{y})^2$$

 $t_0 = min(y)$ 

$$h_0 = \max(y) - t_0$$

$$v_i = \frac{(y_i - t_0)^2}{\sigma_y^2}$$

 $\sum_{i=I}^{N_{int}} u_i \left( x_{k,i} - \overline{x}_k \right)^2$ 

$$\sum_{i=1}^{N_{min}} u_i v_i x_{k,i}$$

$$\sum_{i=1}^{N_{min}} u_i v_i x_{k,i}$$

$$W_{0,k}^2 = rac{\sum\limits_{i=I}^{N_{min}}}{\sum\limits_{i''}^{N_{min}}} W_i \gamma_i \left( \chi_{k,i} - c_{n,k} 
ight)^2$$

### Performance Metrics

$$t_{t} = \frac{\sigma_{x_{t}}}{w_{k}}$$

$$\left| \sum_{i=1}^{N_{t}} u_{i} (\mathcal{J}(\mathbf{x}_{t}) - y_{i})^{2} \right|$$

$$\sum_{i=1}^{N_{r}} u_{i}$$

$$r^{2} = \frac{\left(\sum_{i=1}^{N_{r}} u_{i} (\mathcal{J}(\mathbf{x}_{i}) - \overline{f}(\mathbf{x})) (y_{i} - \overline{y})\right)^{2}}{\sum_{i=1}^{N_{r}} u_{i} (\mathcal{J}(\mathbf{x}_{i}) - \overline{f}(\mathbf{x}))^{2} \sum_{i=1}^{N_{r}} u_{i} (y_{i} - \overline{y})^{2}}$$

$$q^2 = I - s^2 / \sigma_y$$

Descriptor Focus

Standard Error

Residual Error





Figure 4H

401

# Sigmoid Optimization Function

$$f(\mathbf{x}) = t + \frac{h}{1 + \sum_{k=1}^{N_x} e^{-n_k (x_k - c_k)}}$$

$$= s_{y} \left( \sum_{\substack{i=1 \\ \sigma_{y} \sum_{j=1}^{N_{x}} u_{i} \left( f(x_{i}) - y_{i} \right)^{2} \\ \sigma_{y} \sum_{j=1}^{N_{x}} u_{i}} \right) + s_{c} \sum_{k=1}^{N_{x}} \frac{\left( c_{k} - c_{0,k} \right)^{2}}{\sigma_{x_{k}}^{2}}$$

Weighted Mean Squared Error

$$4 \frac{(c_k - c_{0,k})^2}{\sigma_{x}^2}$$
 (

 $S_n \sum_{n} |n_k| \sigma_{x_n}$ 

 $+ s_i \frac{(t-t_0)^2}{\sigma_{iv}^2}$ 

Ih aunbid

#### Sigmoid Optimization Starting Values

$$t_0 = min(y)$$

$$h_0 = max(\mathbf{y}) - t_0$$

$$v_{i} = \frac{\left(y_{i} - t_{0}\right)^{2}}{\sigma_{y}^{2}}$$

$$\left(h_{i} + t_{i} - v_{i}\right)^{2}$$

$$v_i' = \frac{o_{J_i}}{\left(h_0 + t_0 - \mathcal{Y}_i\right)^2}$$

$$\sum_{i,k}^{N_{min}} u_i v_i x_{ki}$$

$$\sum_{i,k} = \frac{\sum_{i=1}^{N_{min}} u_i v_i}{\sum_{i} u_i v_i}$$

 $\sum_{i=1}^{N_{int}} u_i v_i (x_{k,i} - c_{h,k})^2$ 

$$\sum_{i=1}^{N_{inh}} u_i v_i' x_{k,i}$$

$$=\frac{\sum_{i=1}^{I-I}n_{i}}{\sum_{i=1}^{N_{i}}n_{i}v_{i}'}$$

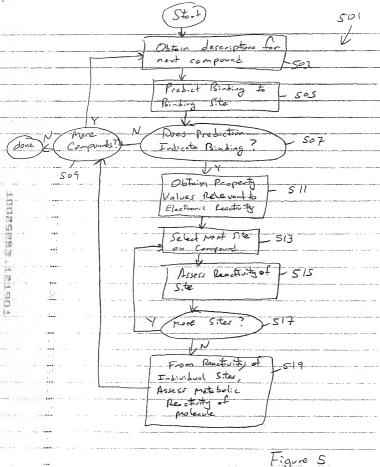
$$\sum_{i=1} u_i v_i$$

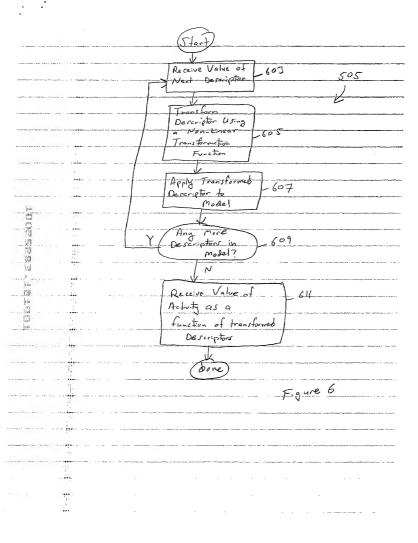
$$C_{i, \ \nu} + C_{i, \ \nu}$$

$$w_{l,k} = \sqrt{\sum_{i=1}^{N_{out}} u_i v_i' (x_{k,i} - c_{l,k})^2} \sqrt{\sum_{i=1}^{N_{out}} u_i v_i'}$$

$$n_{0,k} = \frac{C_{l_1k} - C_{l_1k}}{\sum_{i \in I_{i,k}}}$$

Figure 44





### Optimum logP

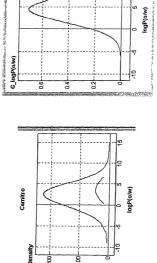
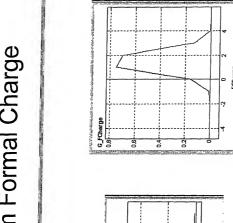
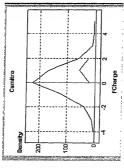


Figure 7A

## Optimum Formal Charge







## Automated Gaussian Fit

3	Jatabase Viewo	y Database Viewer: f./2d6/discrim/01/dac-4.mdb	m/01/d	ac-4.mdb	MONTH SECURITION OF SECURITION	COLUMN STREET,	STATES STATES STATES	COLD TO THE STATE	PLOT SECURED OF SECURED SEC	×	
e luie	Entry, Field Cor	ile Entry, Field Compute District Mindow Yeth	Idow He	, de		ų.				Calle	
-1		A STANDARD S				A MERCHANIST CO.					
	active	MW	PEOE	PEOE_VSA_FHYD	PEOE_VSA_FNEG		SA_FPNEG	PEOE_VSA_FPNEG PEOE_VSA_FPOS	PEOE_VSA_FPPOS	POS *	
		0 130.0780		0.5071	0.3361	361	0.2331	0.6639	0.2	0.2598	
2		0 136.0430		0.5670	0.7154	154	0.4330	0.2846	0.0	0.000	>
e		0 139.1540		0.7499	0.5	0.5756	0.1330	0.4244	0.1	0.1171	
4		0 146.1460		0.5387	0.3	0.3008	0.3008	0.6992	0.0	0.1605	1
9		0 148.2740	12000	W osedele	*( Database Viewer - F 124824isetim/01/paness 4 m/0	discrim/III/e	Apor Fragility	The state of the s	HARMON COLUMNS CONTRACTOR		
9		0 150.2700		Note: Sales	The Fifth Computer Discussion of	T Wordow					
-		0 158.1570	ı.			2 MON.					d T
<b>&amp;</b>		0 160.1800		THE REAL PROPERTY.		NA SALISTING SALES	A STANSSORM WWOTHING	and the second second second	Change and their an an architecture	STATE OF STREET	and the second s
69		1 176.2430	201206	Ē	Field	Center	Width	Height	rmse	fress	12
2	-	0 178.2750	entra e	PROR_VS	PROR VSA PPNEG	0.0311	0.0310	0.8555	0.3764	0.4333	
ξ.	_	1 179.3320	2	FCharge		1.4544	0.5524	0.8927	0.3899	0.3918	0.0002
2	2	0 180.1560	n	PROR VSA PHYD	A_FHYD	0.9182	0.0741	0.7745	0.4176	0.3025	0.4645
5		0 180.2040	4	PROR_VSA-5	3-5	5.1460	5.5009	0.8250	0.4201	0.2941	0.4173
4	-	1 180.2710	G I dálo	PROR VSA-6	.y-6	5.7300	9.7930	0.6950	0.4292	0.2630	0.3797
5	2	1 180.2710	9	a acc		0.5577	2.1359	0.7321	0.4368	0.2368	0.4760
16	9	0 180.3150	200	a base		1.9562	0.8965	0.8122	0.4438	0.2123	0.1308
			60	a_acid		0.000	0.000	0.5993	0.4476	0.1986	0.2228
			60 	logP(o/w)	P	4.4390	2.0997	7 0.7072	0.4484	0.1956	0.1405

Figure 7C

### 2D6 K<sub>i</sub> Model

### Non-linear Size Relation

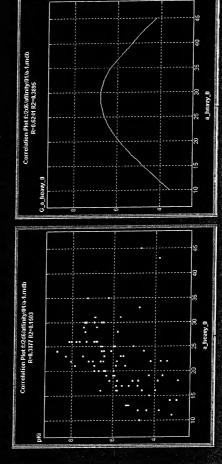


Figure 7D

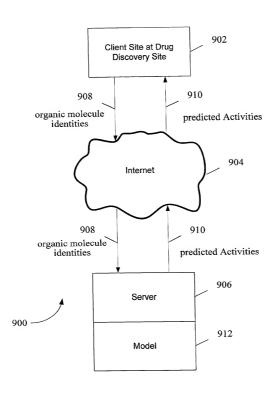


FIGURE 9

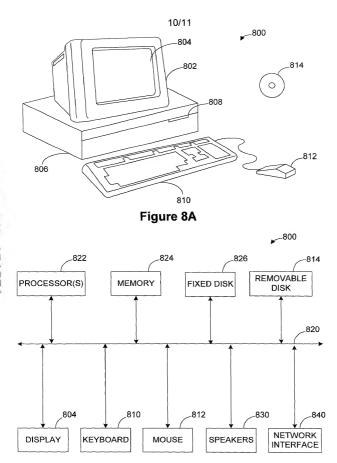


Figure 8B